

Deux remarques sur le traitement formel des textes anciens et ethniques

Christian Bauer

Volume 10, numéro 3, décembre 1977

Sémiotique du discours

URI : <https://id.erudit.org/iderudit/500448ar>

DOI : <https://doi.org/10.7202/500448ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Département des littératures de l'Université Laval

ISSN

0014-214X (imprimé)

1708-9069 (numérique)

[Découvrir la revue](#)

Citer cet article

Bauer, C. (1977). Deux remarques sur le traitement formel des textes anciens et ethniques. *Études littéraires*, 10(3), 501–523. <https://doi.org/10.7202/500448ar>

Tous droits réservés © Département des littératures de l'Université Laval, 1977

Ce document est protégé par la loi sur le droit d'auteur. L'utilisation des services d'Érudit (y compris la reproduction) est assujettie à sa politique d'utilisation que vous pouvez consulter en ligne.

<https://apropos.erudit.org/fr/usagers/politique-dutilisation/>

DEUX REMARQUES SUR LE TRAITEMENT FORMEL DES TEXTES ANCIENS ET ETHNIQUES

christian bauer

Cet article est extrait d'un long manuscrit qui fera, par ailleurs, l'objet d'une série de publications complémentaires : outre les points soulevés dans le présent texte, il y est fait mention, entre autre, d'un certain nombre d'approches sur un traitement formel des textes anciens, permettant une reconstruction au niveau des proto-sociétés et de leur idéologie, notamment dans le domaine Mon-Khmer. De même nous y présentons un certain nombre de travaux orientalistes, sur la métrique par exemple (ceux de Wenk pour le thai, de Hla Pe pour le birman et de J.M. Jacob, F. Huffman et Im Proum sur le Khmer) et sur les contes légaux birmans (de Maung Htin Aung), qui, si on leur portait une plus grande attention, pourraient dans une certaine mesure renouveler les études littéraires occidentales. Nous renvoyons donc le lecteur à des publications futures pour la présentation de ces travaux. (Nous tenons à souligner la collaboration d'Anne Bienjonetti et de Pierre Ouellet pour la version finale (abrégée) du présent article.)

Les méthodes de traitement formel des textes, c'est-à-dire des documents inscrits chacun dans un espace historique spécifique, suffisent-ils à donner un support aux recherches historiques et philologiques ? Voilà la question qui sert de point de départ aux quelques réflexions développées ici. Il s'agit là d'un problème d'une portée épistémologique cruciale : la sémiotique, d'une part, domaine non encore explicitement défini malgré les efforts déployés en vue de l'établir comme science spécifique à côté des autres sciences du langage telle la linguistique, et l'ethnologie, d'autre part, prise au sens d'anthropologie sociale, se trouvent depuis quelque temps dépourvues d'objet; or, des spécialistes de ces deux domaines¹, si étrangers l'un à l'autre, se sont aperçu que leur

¹ Voir, entre autres, Sperber, « Invariants biologiques et universaux culturels » in E. Morin et M. Piattelli-Palmarini (eds), *L'unité de l'homme*, Paris, Seuil, 1972, de même que *Le symbolisme en général*, Paris, Hermann, 1974. Également à consulter les actes du colloque international du C.N.R.S. sur « La situation actuelle de l'anthropologie », organisé par G. Condominas et J. Gamelon-Dreyfus, à Paris, du 22 au 28 avril 1977, à paraître; voir plus particulièrement les interventions d'Ottino et Haudricourt à propos des sociétés à tradition orale et de l'ethnolinguistique.

travail d'exploration et d'exploitation des résultats d'une enquête (des « données brutes ») n'est possible qu'à partir, non seulement d'une représentation, soit métalinguistique, soit d'un type topologique quelconque, mais aussi d'une classification préconstruite de ces données brutes — classification malheureusement trop souvent inconsciente. Depuis quelques années se sont articulées les premières tentatives de classification des données, notamment en linguistique; par ailleurs, le logicien Dana Scott², dans une critique récente où il proposait quelques alternatives en sémiotique formelle, a présenté une liste de critères servant à examiner les préconstruits d'un cadre logique donné. Ce n'est cependant pas le but de cet article de reprendre cette argumentation par ailleurs fort intéressante; nous nous proposons plutôt de présenter, en une première partie, les apories d'un traitement formel de différents types de données et, en une seconde partie, d'examiner des contre-propositions extraites des recherches en cours en ethnologie, préhistoire et linguistique diachronique, plus particulièrement dans le domaine sud-est asiatique et insulindien.

I. Apories

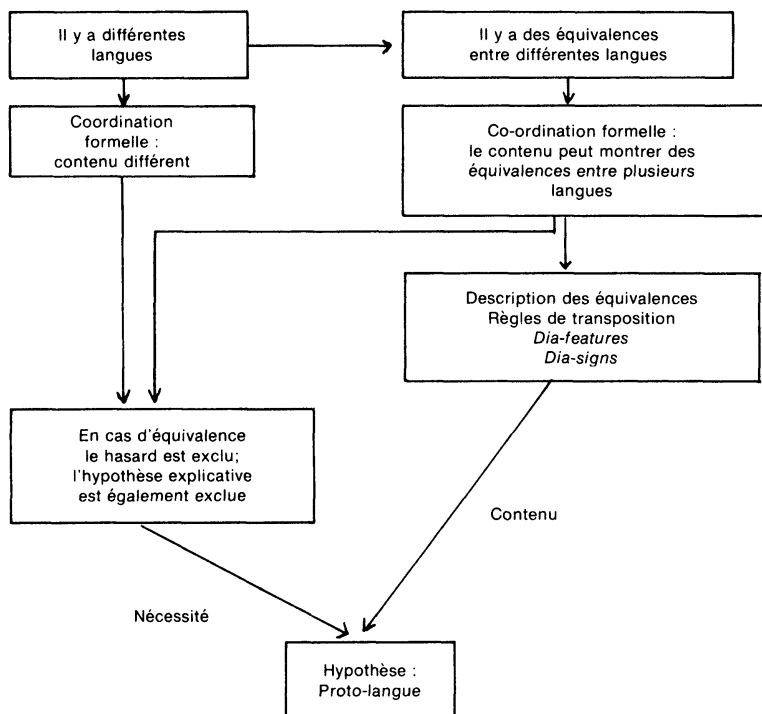
Les quelques apories que nous allons esquisser ici sont à situer évidemment par rapport aux sciences historiques et philologiques, c'est-à-dire aux procédés d'une analyse qui a affaire à des textes sous forme d'inscriptions (épigraphie, paléographie) ou de manuscrits proprement dits ou encore transmis oralement, et de leur intégration dans un cadre diachronique (histoire événementielle). À partir de là, quel que soit le sujet, le domaine et l'hypothèse à élaborer, on essaiera de trouver une base commune — consistant en un certain nombre de traits distinctifs (matrice) ou de différents critères — qui permette une comparaison.

Avant d'entrer dans une discussion sur la classification des données et le problème de leur traitement propre, prenons un

² D. SCOTT, « Background to Formalisation », in H. Leblanc ed., *Truth, Syntax and Modality*, Amsterdam, North-Holland, 1973, pp. 244 à 273.

exemple, simplifié, bien sûr, tiré de la linguistique diachronique³ :

SCHEMA I



ce schéma nous paraît suffisamment explicite et ne nécessite que peu de commentaires; disons ceci : l'argumentation inductive part du fait qu'il y a différentes langues (hypothèse triviale) qui peuvent être, par l'intermédiaire des stades de coordination et de description des équivalences (lexicales, morphologiques) liées génétiquement l'une à l'autre. Nous y

³ Le schéma qui suit, légèrement modifié, est tiré de J. UNTERMANN, « Sprachvergleichende Ausbildung im Grundstudium der Germanistik », in Untermann ed. *Theorie, Methode und Didaktik der historisch-vergleichenden Sprachwissenschaft*, Wiesbaden, Reichert, 1973, p. 79; Untermann présente en plus un ensemble de théorèmes à propos des dérivations diachroniques.

reviendrons plus loin. Ce qui nous intéresse pour le moment, c'est de voir comment une hypothèse, en l'occurrence celle de l'existence d'une proto-langue non attestée dans les documents (inscription, manuscrit), se développe à travers une argumentation fort contraignante, bien que ce ne soit pas l'argumentation elle-même qui mène à l'hypothèse en question, mais plutôt le préconstruit de l'induction. Par ailleurs, en un second temps, nous essaierons de voir à partir de quelles propriétés sémiotiques non encore définies et à quel niveau spécifique s'effectue le travail.

L'argumentation en faveur d'une hypothèse est basée sur une sélection de données, ou, mieux encore, sur une classe particulière de données. On peut, en gros, résumer ces classes dans une liste, qui reste, cependant, provisoire⁴.

A — données en *langue naturelle*

1 — moyen de documentation

a — intuitif

b — manuscrits

i — matériau

ii — datation

iii — trouvailles, lieu (documentation philologique et archéologique proprement dite)

c — transmis oralement

i — conditions de préservation

ii — conditions de performance

d — corpus clos

i — littérature imprimée

ii — moyens de diffusion

B — données dépendant de *paramètres*

C — données dépendant d'*actions*

⁴ Cette liste, modifiée et augmentée, doit son origine à une intervention de Th. Ballmer sur « Empirisme et Linguistique » publié en 1976 dans un volume collectif intitulé *Wissenschaftstheorie der Linguistik* Frankfurt, Athenäum. L'argumentation et les conclusions de Ballmer sont évidemment tout à fait différentes.

D — données à différents degrés de *complexité* (combinaison de plusieurs classes)

Un linguiste, par exemple, qui entreprend des recherches syntaxiques sur des dialectes inconnus essaierait d'abord d'établir une famille paraphrastique d'énoncés simples à partir d'un procédé contrastif (langue première du linguiste, langue première de l'informateur, langue seconde du linguiste et de l'informateur) recourant à la même relation primitive — le travail se fait alors au niveau *A./a.*, c'est-à-dire au niveau des intuitions, sans moyen de contrôle autre que l'appareil formel qui consiste en un certain nombre de règles de correspondance (et d'opérations), construites à partir d'une hiérarchie de relations logiques entre termes primitifs.

Pourtant, ce travail ne nécessite en aucun cas l'élaboration d'une théorie des propriétés sémiotiques quelconques dans la mesure où l'établissement d'une représentation méta-linguistique (métalangue) est déjà dérivé de la logique soit philosophique, soit mathématique; ou encore, on identifiera la sémiotique à une branche spécifique des mathématiques — comme Thom le fait par exemple, du moins implicitement⁵. Rappelons aussi qu'il existe à ce sujet une littérature abondante; J.P. Desclés, dans ce numéro même, examine d'ailleurs la relation entre langage-objet, ou plus précisément une famille paraphrastique dérivée de la même relation dite primitive et méta-langue (ou métalangue).

Ce que nous voulons d'abord montrer, c'est que la classe des données intuitives — en admettant, avec Ballmer, l'inférence qu'il s'agit là d'une véritable classe de données — implique nécessairement le choix d'une argumentation définie par rapport à son *Erkenntnisinteresse*⁶, et en-

⁵ Sebeok, par exemple, considère (communication personnelle, Tampa, Fla., juillet 1975) Thom, avec Peirce, comme sémioticien, se basant, essentiellement, sur son article classique « De l'icône au symbole », repris dans *Modèles mathématiques de la morphogenèse*, Paris, U.G.E. 1974. Signalons que Sebeok conçoit la sémiotique comme basée sur le triplet (1) médecine, (2) logique mathématique et (3) linguistique.

⁶ L'expression française « intérêt de connaissance » est déjà une déformation du terme; cette notion fut adoptée par différentes écoles épistémologiques, notamment en Allemagne et en Grande-Bretagne.

suite, que la sélection d'une classe de données exclut nécessairement un ensemble d'autres procédés analytiques. En ce qui concerne cette dernière affirmation, rappelons tout de suite le travail de Shorto⁷ qui, à plusieurs reprises, a expliqué, démontré et vérifié à partir d'un corpus fort cohérent (inscriptions Mon) que la reconstruction linguistique des stades intermédiaires du développement d'une langue ou d'un groupe linguistique ne peut aboutir à des résultats concrets et corrects qu'à condition que le travail s'effectue au niveau phonologique ou, le cas échéant, au niveau morphologique⁸.

Or, compte tenu du fait qu'il n'y a plus d'informateurs de la langue en question — sauf, éventuellement, pour quelques dialectes archaïques (nous y reviendrons dans la deuxième partie) — il nous manque le moyen d'établir une famille paraphrastique qui nous permette de saisir et de déterminer les relations grammaticales, les hiérarchies syntaxiques et l'ordre des mots (*word order*). Autrement dit — et au risque, de nouveau, d'aboutir à une implication triviale — le corpus disponible (inscriptions, manuscrits sur olles) nous impose l'exclusion de certains types d'argumentation et nous contraint à choisir une seule hypothèse à partir de deux coordina-

⁷ Cf. « Notes on Mon Epigraphy », *BSOAS*, 1958, XVIII, 2, pp. 344 à 352; ou encore ses remarques sur les systèmes d'écriture archaïques, dans *PLS*, tome I. Pour le survol le plus concis sur ce problème (reconstruction linguistique) nous renvoyons à H.M. Hoenigswald, *Language Change and Linguistic Reconstruction*, University of Chicago Press, 1960.

⁸ Évidemment, nous connaissons, à partir d'un ou plusieurs lexiques, diverses classes de mots (noms, verbes, attributifs, etc.); autrement, une traduction, (que nous avons déjà, dans la plupart des cas), serait impossible. Mais ici, il s'agit seulement de l'impossibilité d'établir des paraphrases (contrôlables) à partir des relations logiques et des règles de dérivation syntaxique. Rappelons que certains linguistes, notamment aux États-Unis, se voient confrontés à ces questions en travaillant sur les aspects diachroniques de la syntaxe. Mais, par contre, il est symptomatique que quelqu'un comme Haudricourt consacre tout son intérêt uniquement à la phonologie diachronique (voir son recueil *Problèmes de phonologie diachronique*, Paris, SELAP, 1973). On peut établir une série de paraphrases, bien sûr, le matériau des inscriptions est suffisamment riche, mais le procédé et l'argumentation seront nécessairement différents d'un travail synchronique qui se base sur les intuitions d'un informateur, et, comme nous l'avons dit plus haut, sur une stratégie contrastive. Voir également Haudricourt, « Why and How the « Small Languages » should be studied », *JSS*, 1975, LXIII : 2, pp. 1 à 4.

tions formelles comme nous l'avons esquissé dans le schéma 1.

L'autre point d'intérêt de ce problème concerne l'enjeu sémiotique proprement dit, à partir de la supposition peut-être classique⁹, qu'une théorie linguistique se clôt sur un autre domaine qu'une théorie sémiotique.

L'intérêt d'une théorie linguistique est double : ou bien on reconstruit un ensemble de proto-langues non attestées dans les documents à partir des dialectes archaïques encore existant ou du matériau (fragmentaire) disponible — l'intérêt est alors *historique* (et non seulement diachronique) puisque les diverses proto-langues peuvent se regrouper géographiquement et indiquer des migrations¹⁰ — ou bien on cherche à travers soit des propriétés sémantiques soit des unités syntaxiques d'une langue particulière ou d'un ensemble de langues (apparentées ou non peu importe)¹¹ des con-

⁹ À consulter et à retravailler, les travaux de Carnap, leur revision par H.H. Lieb récemment, et, depuis quelque temps, les remarques originales de Halliday (1973) et de Lyons dans une série de livres à paraître sur la sémantique.

¹⁰ Cf. l'article de I. Dyen, malheureusement ancien, « Language Distribution and Migration Theory », in *Language*, XXXII, 4, 1956, pp. 611 à 626. Pour l'aire tibéto-birmane et mon, voir (presque) tous les travaux de G.H. Luce, entre autres dans *JBRs*, XLII, 1, 1959, et son œuvre capitale *Old Burma — Early Pagan*, 3 vol., New York/Ascona, Artibus Asiae, 1971. Luce utilise la linguistique diachronique (ou encore comparative) comme moyen de reconstruction pré- et proto-historique; selon sa stratégie il examine les résultats linguistiques par rapport aux trouvailles archéologiques, paléographiques liées à l'architecture. Un procédé pareil est indiqué également par Haudricourt, dans « Note de géographie linguistique austroasiatique » in Ba Shin *et al.*, eds., *Essays offered to G.H. Luce*, Artibus Asiae, Supplementum, XXIII, 2 vol., Vol. 1, 1966, pp. 131 à 138.

¹¹ Nous laissons de côté les questions concernant la linguistique descriptive (analyse qui se fait sur le terrain) ou la grammaire d'une langue particulière (dans la terminologie anglo-saxonne *Reference Grammar*). Pour les points mentionnés ici, voir aussi H. Schnelle, « Foundations of Theoretical Linguistics », in *Logic, Methodology and Philosophy of Science*, IV, (1971), Amsterdam, North-Holland, 1975. Ailleurs Schnelle (cf. note 4 de notre article) esquisse une théorie des *Sprachausschnitte* (fragments linguistiques, ou, pour être plus proche de la métaphore, tranches linguistiques d'un corpus) qui coïncide en partie avec notre classification des données, (cf. plus haut), mais qui rend encore compte des langues artificielles et spécifiques (*Fachsprachen*) et d'autres implications socio-linguistiques et pragmatiques. Une telle discussion nous mènerait trop loin.

traintes universelles et des opérations généralisables — en quel cas, la théorie linguistique se heurte à la réalité cognitive, à savoir à des questions d'ordre psychologique et neurologique qui en constituent la clôture. L'alternative — et la seule — qui se pose pour le linguiste est la suivante : ou bien il se situe dans une perspective, soit diachronique, soit descriptive, ou bien il est obligé, dès le départ, de construire une théorie à la fois des propriétés formelles de son cadre explicatif et de l'apprentissage du langage, donc, de se poser le problème du statut des universaux.

On peut alors exprimer l'aporie sémiotique, dont nous parlions plus haut, en ces termes : excluant, (non par apriorisme, mais pour des raisons de clarté dans l'argumentation) l'hypothèse d'une identification de la sémiotique à une branche spécifique des mathématiques, nous pourrions considérer que la portée du champ sémiotique peut s'étendre à la modélisation de différentes formes d'organisation, soit cognitives (Pierre Maranda parle de *semantic charts*), soit sociales, soit, encore, linguistiques, et tomber alors, à nouveau, dans le piège épistémologique qui consiste en ceci que l'élaboration première d'un procédé heuristique et d'un appareil formel susceptible de saturer en quelque sorte l'organisation des données — soit, l'imposition d'une forme d'organisation, même de type topologique très tolérante et flexible, à un ensemble de données — n'aboutit le plus souvent qu'à des généralités telles qu'il est impossible pour un chercheur travaillant au niveau des particularités — par exemple des catégories indigènes (en ethnoscience) — d'y trouver un quelconque intérêt : dans ce cas, le champ sémiotique rejoint purement et simplement celui d'une philosophie des sciences.

Par contre, l'examen (au sens de *test*) d'un procédé analytique appliqué à un phénomène particulier, mais voisin quand même d'autres phénomènes (apparentés), se situe immédiatement dans la méthodologie du domaine concerné.

Prenons deux exemples pour clarifier ce point de vue : le premier se référant à la linguistique comparative (diachronique), le second à l'origine de l'art indo-javanais. En ce qui concerne le premier exemple, considérons d'abord cet

énoncé de R.M.W. Dixon¹² à propos des classifications régionales et génétiques des langues australiennes :

It is improper to draw an immediate inference of (possible or probable) genetic relationship from an observation « both language A and language B have feature X but many surrounding languages lack this », whether the X be shared vocabulary, or some phonological, grammatical or semantic property (. . .). We are thus saying that possible positive evidence from each of the languages being compared — for example A and D have property X, but B and C show Y in its place — is more likely to be genetically significant than merely negative evidence — A and D have X but B and C lack it. The latter case is more liable to be due to a simpler diffusion of « a tendency to eliminate X ».

Il s'agit ici de l'examen d'une inférence ambigue à l'intérieur d'une coordination formelle (voir le schéma 1) et à partir d'une implication qui paraît, à première vue, simple. Toutefois, il n'en est pas ainsi; car nous avons affaire à deux séries de propositions différentes, à savoir¹³ :

- (1) $L(A) \wedge L(B) \rightarrow X.$
- (2) $L(P_1 \dots P_k) \vee L(Q_1 \dots Q_n) \rightarrow \sim X.$
- (3) $L(A) \wedge L(D) \rightarrow X.$
- (4) $L(B) \vee L(C) \rightarrow Y.$
- (5) $((L(A) \wedge L(D) \rightarrow X) \wedge (L(B) \wedge L(C) \rightarrow \sim X))$
 $\mapsto (\underset{q}{\overset{p}{F}} (X) \longleftrightarrow L(p.A \wedge D) \wedge L(q.B \wedge C)).$

¹² R.M.W. DIXON, « Tribes, Languages and other Boundaries in northeast Queensland » in *TBA*, p. 230-231.

¹³ Les symboles utilisés ici se lisent : « < » conjonction (et), « > » disjonction (ou), « \rightarrow » implique, « \longleftrightarrow » s'impliquent réciproquement, « \sim » négation (*non X*), « F » sous-ensemble flou (dans le cas de la proposition (5) se lit « sous-ensemble flou F s'étend sur l'espace de tolérance (ou son intégral) de q à p »). Les lettres « L », « X » se lisent respectivement « langue (particulière) » et « propriété » (ou « trait », en anglais *feature*); « A », « B », « C », « P » et « Q » représentent des langues particulières, tandis que les indices « i », « k », « j » et « n » signifient, appliqués ici à leur contexte précis, une variété; autrement dit, les deux expressions « $L(P_1 \dots P_k)$ » et « $L(q_1 \dots Q_n)$ » donnent une variété (celle de i à k et de j à n respectivement) de deux langues particulières, « P » et « Q ». Cette variété, au sens topologique, est dans notre cas une classification régionale (*areal topology*, ce que Dixon dénomme, par commodité, *surrounding languages*). Rappelons que le symbole « \rightarrow » indique, grossièrement, l'application de deux propositions à un espace de tolérance.

La première série est constituée des propositions (1), (2) et (5), la deuxième de (3) et (4). La première marque, formellement, les paraphrases respectives :

(1a) Les deux langues *A* et *B* ont la propriété *X*.

(2a) Les deux ensembles des langues *P* ou *Q* dont les indices représentent une variété régionale n'ont pas la propriété *X*.

(5a) Les langues *A* et *D* ayant la propriété *X* et les langues *B* et *C* ne l'ayant pas forment un sous-ensemble flou *F* sur l'intervalle de *q* à *p* partageant la même propriété *plus ou moins* (facteurs de turbulence, de diffusion, etc.). Si la propriété *X* figure dans cet espace de tolérance (de *q* à *p*) et seulement si, alors les langues *A* et *D* se trouvent dans un voisinage de *p* et *B* et *C* dans celui de *q*.

Avant d'interpréter ces paraphrases, présentons la deuxième série :

(3a) Les langues *A* et *D* ont la propriété *X*.

(4a) Soit la langue *B* soit la langue *C* a la propriété *Y*.

Rappelons, en passant, que nous n'utilisons pas ici une notation liée aux problèmes de quantification à partir de laquelle — selon les travaux de Hintikka, par exemple, — chaque série de paraphrases pourrait être intriquée l'une dans l'autre par un système de quantification ramifiante (*branching quantification*). Ce problème est, pour le logicien, passionnant; mais pour des raisons de clarté, nous l'écartérons ici.

Si nous interprétons maintenant cette démarche, nous dirons que nous effectuons d'abord un premier pas, intuitif, d'observation sur le terrain, puis, une sorte de pas-de-deux logique. Ce premier pas, motivé par un intérêt strictement linguistique — et non sémiotique —, cherche à lier plusieurs phénomènes observables partageant certains traits généralisables (issus souvent d'un cadre théorique lié, lui aussi, à l'intérêt du chercheur). À partir de là, comme le montre de nouveau notre *schéma 1*, on commence à postuler une hypothèse explicative — motivant le pas-de-deux logique, qui explicite ce que Untermann appelle « co-ordination formelle » pour désigner la transposition du problème au niveau des propriétés formelles. Dès lors, l'intérêt ne porte plus que sur la

validation des relations logiques en jeu, sur l'élaboration d'un cadre formel et sur la relation entre un ensemble de paraphrases — (1a) à (5a) — et les propositions logiques — (1) à (5).

On ne concentre plus son intérêt sur la question de l'existence supposée d'un ensemble de proto-langues et de relations génétiques quelconques entre différentes langues parlées dans différentes régions, mais plutôt sur le problème de l'élimination des contradictions et des paradoxes dans un modèle mathématique (au sens strict du terme). Le linguiste, quant à lui, ne fait qu'un pas vers l'abstraction afin d'en tirer un certain nombre de règles dérivationnelles, comme dans le cas suivant¹⁴ :

- (6) a. djinaŋ
 gunaŋ
 b. djina
 guna

où dans la colonne *a.* le *ŋ* marque la finale des noms-racines (*root nouns*) tandis que les exemples de la colonne *b.* montrent l'absence d'une telle finale; d'autre part, les langues Djabugay et Yidinj ajoutent à leurs noms-racines le morphème *-ŋ* qui représente le présent et le non-passé, respectivement.

À partir de cette observation simple, — présence ou absence d'un phonème, soit final soit initial, existence d'un morphème de valeur aspectuelle, dans d'autres langues, susceptibles cependant d'être apparentées aux premières, ou même l'existence d'un ergatif, tel en Yidinj *-ŋgu* (qui se réduit à *-ŋ* dans le cas des racines disyllabiques) — on construira une hypothèse concernant les règles dérivationnelles pour différentes langues, règles qui permettent alors une comparaison. Le linguiste cependant n'a pas besoin de développer son formalisme jusqu'à inclure l'appareil formel ((1) à (5)) mentionne plus haut, dont le rôle est en fait de niveler les différences importantes, nécessaires à la classification du problème des relations génétiques¹⁵. Revenons au partage

¹⁴ Cf. Dixon, *op. cit.*, p. 222.

¹⁵ Encore une fois, le problème se pose autrement dans la syntaxe, où l'on a, en effet, besoin des règles opérationnelles pour contrôler ces paraphrases.

que nous avons fait de ces propositions en deux séries : (1), (2) et (5) donnent respectivement des règles pour la présence d'un trait commun à deux langues (conjonctif, règle (1)), pour l'absence d'un trait commun à un ensemble de langues (disjonctif, règle (2)), et pour la projection (ou l'application, *Abbildung*) des règles (1) et (2) sur un espace de tolérance (règle (5)) dont les valeurs maximales et minimales, garantissant la stabilité structurelle de l'espace, se situent entre le voisinage de deux langues possédant le trait X et celui qui marque l'absence de X dans d'autres langues. Cette série de propositions, qu'elle pose le problème en termes discrets ou continus (par un espace de tolérance ou, autrement dit un ensemble de sous-ensembles flous) le pose toujours sur le plan quantitatif, selon le couple minimum-maximum. Toutefois, pour le linguiste, la deuxième série de propositions — les formules (3) et (4) — représente, comme nous l'avons vu dans la citation de Dixon, une valeur heuristique beaucoup plus solide, par le fait qu'il s'agit là, en effet, d'un changement qualitatif, au sens où la proposition (3) marque le trait X tandis que la proposition (4) — qui est disjonctive — représente le remplacement de ce trait par Y. Laissons pour le moment cette réflexion en suspens, nous y reviendrons plus loin.

Le deuxième exemple est tiré d'un domaine tout autre, celui de l'art indo-javanais et de son développement. Depuis le début du siècle, deux théories sur l'origine de l'art indo-javanais rivalisent en un débat interminable, celle de Parmentier¹⁶ — architecte de formation — et celle de Bosch — philologue et archéologue formé dans la tradition néerlandaise¹⁷.

Parmentier part d'une analyse formelle des éléments *isolés* des structures architecturales, comme la morphologie des arcs, des niches, l'emplacement des portes, etc., et fait l'hypothèse, sur cette base, d'une influence directe de l'art cham et khmer sur l'architecture javanaise; sa conclusion se fonde donc uniquement sur des éléments minimaux

¹⁶ H. Parmentier « L'architecture interprétée dans les bas-reliefs de Java » *BEFEO*, VII:1, 1907, pp. 1 à 61.

¹⁷ F.D.K. Bosch, « A Hypothesis as to the Origin of Indo-javanese Art », *Rupam*, 17, 1924, pp. 6 à 41.

morphologiquement identiques, et fait abstraction du rôle que joue dans son interprétation son propre savoir le fait, par exemple, que l'art cham et khmer ont déjà été suffisamment analysés et sont en quelque sorte implantés dans un contexte scientifique approprié. La première étape de son argumentation s'articule sur deux réduction de différents niveaux : d'abord il écarte toutes les questions d'ordre linguistique, à savoir le statut, dans les civilisations respectives, de la littérature ethnique, transmise oralement, des textes écrits, (de caractère religieux, légal, philosophique, littéraire), et les différentes pratiques sociales (système parental, pratiques rituelles, implications économiques, migrations inter-ethniques, position des pouvoirs, des monastères), bref, le corps social, ou mieux encore, pour reprendre la belle métaphore du traducteur allemand de Toynbee, Kempfski, *Gesellschaftskörper* (chez Toynbee lui-même, simplement *society*); ensuite, une deuxième réduction se manifeste dans son cadre d'analyse même : la morphologie globale d'une structure architecturale quelconque est décomposée (*split off*), réduite donc à des éléments minimaux ou isolés, tel qu'indiqué plus haut. C'est là, nous semble-t-il, l'évidence même de la fausseté de cette conjecture — selon laquelle l'art javanais serait directement influencé, et, par conséquent, historiquement postérieur à l'art cham ou khmer, —; puisque une étude morphologique qui ne rend pas compte de l'assemblage (du *setting*), de l'agencement, des éléments en question, n'a aucune valeur heuristique.

Bosch, par contre, part d'une analyse globale des temples de l'Inde du sud et de Java, et constate qu'il n'y a pas d'identité globale, mais seulement des éléments minimes semblables — relevant effectivement de lois statiques. De là, il lance une première hypothèse selon laquelle, l'influence, dans le cas où il y en aurait eu une, ne serait pas directe, mais médiatisée. La première justification de son hypothèse est que s'il est possible (bien qu'invraisemblable) qu'il y ait eu une époque où des architectes indiens venaient à Java pour construire des temples hindous à l'aide d'esclaves javanais, ces indiens, lettrés, auraient laissé des traces écrites, en sanskrit, dans diverses sortes de documents, notamment des inscriptions; or, on ne trouve aucune trace de ces inscriptions. Une autre possibilité est que ces immigrants, n'ayant guère de

connaissance du sanscrit, parlaient plutôt des dialectes de l'Inde du sud, comme le Tamil, le Telugu, etc., mais alors ils ont dû être, linguistiquement et socialement, fortement isolés dans la communauté javanaise, et, par conséquent, ils n'ont dû exercer aucun pouvoir politique et économique sur les javanais. Bosch renverse alors l'hypothèse de la façon suivante : si ce ne sont pas des indiens du sud qui construisaient ces temples, il faut que les javanais aient eu connaissance de leur architecture. Conjecture probable, et, en fait, seule explicative : à partir *du* texte.

C'est un fait bien connu des Indo-germanistes qu'il existe un certain nombre de traités d'architecture, sans doute auparavant intégrés dans des textes religieux, appelés *śilpāsāstras*. Une étude approfondie de ces *śilpāsāstras* montre une dénaturation dans la grammaire et le style sanscrit; Bosch : « The later the Sastra the more corrupt is the sanskrit. » Deuxième conjecture : il y a alors des variations dialectales dans plusieurs versions non originaires de l'Inde qui relèvent d'une documentation écrite hybride, à savoir que certains *sāstras* circulaient fort probablement à Java et que d'autres étaient transmis oralement et dictés ensuite.

Mais regardons ce problème de près : ces *śilpāsāstras*, toujours liés à la pratique religieuse et au système philosophique, montrent, entre autres, un système numérique de mesures corporelles et de mesures d'architecture appliqué à une nomenclature fort précise.

L'analyse de Bosch explicite à ce stade, à partir de nombreux exemples, que les javanais appliquaient *exactement* ce système numérique, mais toujours en gardant une certaine liberté, pour les éléments d'ornement, par exemple. Somme toute : premièrement, les javanais eux-mêmes construisaient les temples, deuxièmement la construction suit une forme canonique dictée par des traités d'origine religieuse (de l'Inde), troisièmement le style proprement javanais fut développé à partir des éléments qui ne sont pas liés à la forme canonique.

Quelle est alors la conclusion de la démonstration de ces deux théories différentes.

Réduire une structure globale à des éléments morphologiques isolés et exclure l'environnement fonctionnel qui

les détermine — soit formellement, (interne), soit historiquement ou socialement (externe) — impliquent *nécessairement* une déformation des données, quel que soit le niveau d'abstraction. Le choix des éléments à comparer devient entièrement arbitraire, comme le démontre l'exemple de Parmentier, et mène à des conclusions aberrantes. Par contre l'approche de Bosch intègre plusieurs composantes à partir *du* texte — documentation écrite, orale, — et, par là, dégage non seulement des déterminants historiques mais aussi le statut de différentes pratiques sociales. Une étude dite sémiotique explorant le niveau morphologique des structures architecturales n'aboutira à aucun résultat concret; mais en comparant plusieurs pratiques rituelles, les prescriptions légales et religieuses, les relations économiques et politiques des communautés en question, bref un ensemble de faits sociaux irréductibles à des réseaux (*networks*) logiques, on retrouvera sans doute des facteurs déterminant la morphologie à étudier (architecture, sculpture, bas-reliefs).

En guise de conclusion, nous pouvons constater, en gros, l'existence de trois identifications possibles d'une théorie dite sémiotique, et cela nous ramène à l'aporie dont nous avons parlé plus haut, celle qui veut que la sémiotique s'absorbe *toujours* dans une branche spécifique d'un domaine spécifique :

Proposition 1 : On identifiera la *sémiotique* à l'étude formelle des morphologies quelconques et au problème de leur modélisation; on est alors forcé de l'intégrer ou bien dans des méta-mathématiques (théorie des modèles, théorie des catégories) ou bien dans la topologie soit classique soit de type Thom-Zeeman.

Proposition 2 : On identifiera la *sémiotique* à l'étude :

- a. de la classification des données
- b. de l'examen de l'argumentation
- c. du contrôle des procédés d'analyse
- d. de la construction des systèmes modélisants; et on l'intégrera dans une épistémologie et une méthodologie générales (Peirce) faisant partie de la philosophie des sciences (*Wissenschaftstheorie*).

Proposition 3 : On identifiera la *sémiotique* à l'étude des universaux et on l'intégrera dans la psychologie cognitive.

En ce sens, la citation suivante de R. Needham, éminent spécialiste des systèmes parentaux, nous paraît exemplaire et plus explicite que la critique cavalière de Haudricourt qui considère toute approche formelle méchamment comme de la métaphysique (non sans raison) :

Après un examen minutieux¹⁸, nous parvinmes à la conclusion que les techniques mathématiques étaient inadéquates à l'étude de sociétés aussi simples et closes que les systèmes prescriptifs, et que les modèles de permutation n'étaient ni plus commodes ni plus précis que les procédés mathématiques (...). Je ne voudrais pas cependant que l'on voie dans ces réserves une quelconque hostilité de principe vis-à-vis de l'analyse formelle (...). Ce n'est qu'une question de méthode. Que l'on soit déçu par l'échec de techniques importées de la linguistique, n'implique pas que l'on rejette par principe tout analogue ou toute impulsion venant de cette discipline. (...) De même, objecter, comme je le ferai, que les prétendues analyses transformationnelles de terminologies de la parenté portent mal leur nom — pour la bonne raison qu'elles n'opèrent aucune transformation — ne signifie pas qu'on nie la possibilité de s'inspirer des parallèles linguistiques sur le plan théorique¹⁹.

Ce que Needham applique ici à l'ethnologie qui emprunte une base méthodologique à la linguistique — comme le structuralisme depuis 1949 — est également valable pour les chercheurs travaillant sur des textes anciens et ethniques utilisant des techniques de modélisations mathématiques. Quelques contre-propositions pour une étude formelle seront données dans la deuxième partie.

II. Contre-propositions

Dans cette deuxième partie nous allons essayer de proposer, par une présentation *positive*, quelques travaux en cours sur une approche formelle des textes ethniques et anciens; nous ne considérerons ici, par manque d'espace, qu'un type d'approche que nous appellerons approche génétique, par rapport aux approches thématiques et régionales (représentées, entre autre, par les travaux d'Ottino et Vogel et

¹⁸ Il fait référence ici à F. Korn et R. Needham, « Permutation models and Prescriptive Systems » in *Man*, V, 1970, pp. 393 à 420.

¹⁹ Tiré de *La parenté en question*, Paris, Seuil, p. 31 et 32.

ceux de Denise Paulme²⁰) et une approche de type fonctionnelle comme celle de G. Dumézil — nous renvoyons le lecteur à des publications ultérieures pour l'analyse de ces deux dernières approches.

L'approche génétique dans l'étude des grands textes épiques a pour objectif principal d'en retracer les nombreuses variantes jusqu'à leur origine, si possible, et de développer, à partir de là, une théorie des regroupements géographiques. Ce qui la distingue, par contre, des approches que nous appelons fonctionnelles (Dumézil), c'est l'absence d'une reconstruction des proto-types, dans ce cas, épiques. Notre utilisation du terme *génétique* est donc malheureusement légèrement ambiguë.

Nous choisissons deux auteurs, R.A. Stein, tibétologue et sinologue au Collège de France, et J. Garanger, archéologue spécialiste des Nouvelles-Hébrides, qui, tous deux, malgré l'apparente divergence géographique et analytique, l'un étant philologue, l'autre pré-historien, ont tenté un examen des sources orales liées aux pratiques rituelles et aux fonctions sociales en général, par rapport aux sources historiques soit sous forme écrite soit par le moyen de fouilles.

Présentons d'abord les recherches de Stein²¹; rétablissant en quelque sorte le schéma de son approche, nous pouvons constater sept repères principaux :

1. Inventaire;
2. Pays et nom du héros;
3. Modélisations des thèmes;
4. Éléments constitutifs de l'épopée et du héros;
5. Sources;
6. Linguistique;
7. Étude du héros.

²⁰ Pour les travaux de P. Ottino et Cl. Vogel, voir « La reproduction du merveilleux dans le conte malgache », *ASEMI*, 1976, VII:2-3, 137-163; pour ceux de D. Paulme, *La mère dévorante- Morphologie du conte africain*, Paris, Gallimard, 1976.

²¹ Cf. R.A. STEIN, *Recherches sur l'épopée et le barde au Tibet*, Paris, P.U.F., 1959.

Interprétons maintenant chacun de ces points de repère : Après avoir établi une liste des différents épisodes, Stein commence à recueillir toutes les sources écrites, à savoir les manuscrits et les xylographes. À partir de là, il s'occupe des traces *localisables* qui consistent en des éléments de classement, de datation et en la filiation des versions — versions indigènes, versions étrangères, versions écrites, versions orales recueillies récemment, mais probablement assez anciennes. Ensuite, comme deuxième repère, il étudie le pays et le nom du héros en essayant de retrouver les noms-de-lieux et les noms de tribus mentionnés dans l'épopée et ses épisodes sur des cartes, soit anciennes, soit récentes, à la suite de quoi il fait une étude historique des réminiscences proto-historiques. Le troisième repère est consacré à la modélisation des thèmes; l'auteur propose une classification des fonctions quadripartites et tripartites du héros des fragments épiques anciens et des thèmes apparemment déplacés. Signalons, en passant, qu'il s'agit là d'une différence fondamentale d'avec le structuralisme classique, dans la mesure où certains de ses adeptes, notamment en ethnologie de l'Amérique du Sud, refusent l'existence des proto-types mythiques *historiques*, ou, pour être plus précis, l'existence d'une version première ou encore, ce qui nous semble plus probable après des décennies de débat chez les Indogermanistes, un ensemble certain de versions premières liées en un faisceau de fonctions idéologiques communes. En ce qui concerne la question du *formalisable*, nous avouons, à nouveau, notre désintérêt; car une formalisation des thèmes basée sur une matrice simple recourrant à des traits distinctifs — qui, rappelons-le, sont difficiles à déterminer selon leur pertinence — n'aboutit à rien d'autre qu'à elle-même et n'a aucune valeur heuristique, puisque la variation — et c'est elle qui nous intéresse — se trouve dans les noms-de-lieux, les noms propres et les accessoires mentionnés dans les différentes versions. Réduire donc les thèmes à un réseau formel des propriétés *sémiotiques* — ce qui pose, d'ailleurs, une fois de plus, le problème de la différence entre les propriétés sémiotiques et les propriétés *sémantiques*²² — revient à effacer (*a priori* !) les variantes.

²² Voir SPERBER, *op. cit.* et BENVENISTE, « Sémiologie de la langue », in *Problèmes de linguistique générale*, II, Paris, Gallimard, 1974, 43-66.

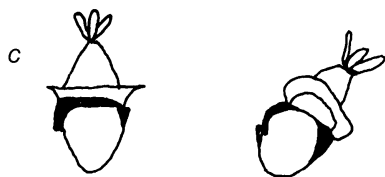
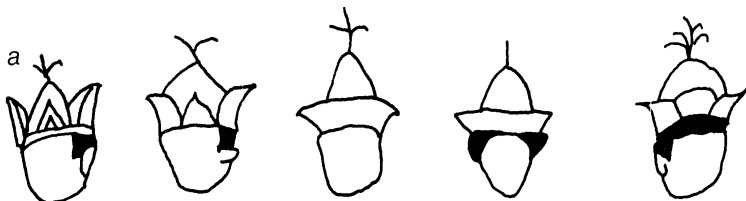
Les éléments constitutifs de l'épopée et du héros forme le quatrième repère; dans cette partie, Stein examine les différentes fonctions sociales de la récitation, les affinités du barde (entraînement; milieu social, statut) et les compare avec des techniques chamaniques. Lié étroitement au repère précédent, le cinquième est consacré aux sources sociologiques et religieuses; après avoir fait une distinction entre tradition populaire et moderne, l'auteur nous présente l'adaptation bouddhique des fêtes, légendes et énigmes, et aussi des chants généalogiques. Lorsqu'il s'agit d'une tradition très ancienne, la deuxième partie décrit aussi des joutes oratoires, des danses, des rituels et des cultes conservés jusqu'à aujourd'hui.

Le sixième repère de la démarche nous paraît plus proche d'un intérêt sémiotique proprement dit; ici toutes les questions d'ordre stylistique sont discutées. Mais, à nouveau, nous avons du mal à voir comment des spécialistes de sémiotique littéraire peuvent pousser l'analyse du texte au-delà de celle des linguistes, puisque la prosodie, l'étude des épithètes, des métaphores et des formules épiques (au sens de Parry) est depuis longtemps un domaine bien établi chez les philologues. Bien sûr, E.R. Haymes, dans une étude récente²³, propose d'étudier à fond, enfin, les épithètes épiques dans les versions orales et d'utiliser des techniques descriptives empruntées aux théories transformationnelles ou relationnelles, d'établir des hiérarchies de contrainte syntaxiques qui excluent l'emploi de telle ou telle formule épique. Mais, comme nous l'avons toujours dit, dès qu'on se lance à un niveau d'abstraction des données qui permet une formalisation, le cadre formel ainsi élaboré perd tout de suite son intérêt pour les chercheurs qui s'occupent d'une région spécifique. Encore une fois, l'étude *génétique* (et régionale) exclue toute formalisation possible parce que les variantes ne sont pas réductibles.

Pour conclure cette courte présentation, discutons encore de la question d'une approche sémiotique possible du cadre récitatif de l'épopée de Gesar étudiée par Stein. Ce que nous allons dire peut paraître, à certains lecteurs, anecdotique et

²³ Cf. E.R. HAYMES, *Das mündliche Epos*, Stuttgart, Metzler, 1977.

plutôt ironique; il s'agit d'un objet : le chapeau. Comme le montrent les illustrations 1 et 2, il s'agit de deux séries de chapeaux du barde tibétain, retracées par nous de l'original²⁴.



Après avoir examiné les différentes formes de récitation de l'épopée de Gesar, Stein étudie les accessoires (et ensuite encore, l'origine, la formation, etc.) du barde tibétain. Il établit entre autres, une typologie des chapeaux. La raison en est très simple : dans différentes régions, dans différentes tribus et — ce qui n'est pas prouvé, mais bien probable — à différentes occasions, le barde porte un chapeau de morphologie différente, que ce soit la pointe qui diffère, ou la « cloche », ou la bande liée directement à la tête du barde et couvrant, en partie, ses cheveux.

²⁴ STEIN, *op. cit.*, p. 344, colonnes a, b et c.

Selon l'approche génétique, l'identification de variantes (ou d'une version autre, ou même d'un thème déplacé dans la proto-version) trouve un support considérable dans l'analyse des accessoires du barde. Par exemple, on peut facilement établir une *homologie*²⁵ (au sens strict) entre la morphologie du chapeau (en tant qu'accessoire parmi d'autres) et les variantes et leur *voisinage*, compte tenu du fait qu'il y a aussi (à degrés variables) voisinage dans les variantes différentes de l'épopée et de ses épisodes. Un sémioticien peut faire une étude comparative du chapeau des bardes dans différentes aires linguistiques et culturelles, comme, par exemple, entre la Laponie et le Tibet, mais cette étude, nous le soulignons à nouveau, se heurtera plutôt à des contraintes physico-anatomiques qu'à des contraintes culturelles. Ici, chez Stein, il s'agit d'un moyen d'identification, donc d'un moyen heuristique, qui permet, le cas échéant, une falsification de l'hypothèse, ou encore un support vérifiable. L'intérêt principal porte alors sur les relations génétiques à partir desquelles on établira les hypothèses proto-historiques de différentes organisations sociales originaires.

Le deuxième travail de cette catégorie, celui de J. Garanger²⁶, relève d'un tout autre domaine, à la fois géographique et scientifique, à savoir l'archéologie des Nouvelles Hébrides. Après avoir constaté l'échec des travaux des Européens au XIX^e siècle sur la valeur historique des traditions orales en Polynésie, l'auteur par le moyen de fouilles archéologiques examine la véracité de deux grands cycles mythiques néo-hébridais (mélanésiens), celui de *Kuwae* et de *Roy Mata*.

Le premier cycle mythique parle d'une île légendaire aux îles Shepherd et nomme une communauté détruite par une éruption volcanique. Le mythe donne en effet quelques expli-

²⁵ Il nous paraît intéressant d'attirer l'attention sur *Tristes tropiques* de Lévi-Strauss (bien que nous ne discutons pas ici son approche) qui établit également une homologie entre les variations des tatouages des Indiens de l'Amérique du Sud et leur organisation sociale spécifique.

²⁶ J. GARANGER, *Archéologie des Nouvelles Hébrides*, O.R.S.T.O.M., Publications de la société des Océanistes, Paris, Musée de l'Homme, XXX, 1972; et « Tradition orale et préhistoire en Océanie », *Cahiers O.R.S.T.O.M.*, série sciences humaines, XIII:2, 1976, pp. 147 à 161.

cations sur le déroulement de la catastrophe ainsi que sur le héros. Le premier pas fut de vérifier cet événement par des recherches géologiques (qui se sont révélées positives) et, ensuite, par des fouilles archéologiques qui permettent une datation définitive de la poterie et de l'outillage (600 av JC). En plus, la localisation de l'organisation des sépultures sont conformes aux données de la tradition. Par ces moyens, on a donc obtenu la certitude de l'ancienne existence de l'île légendaire détruite par un cataclysme, de la date de la catastrophe et de la réalité du héros mythique.

Le deuxième cycle mythique parle d'un autre héros mythique, *Roy Mata*; des fouilles furent également entreprises en un lieu désigné par le récit et ont abouti à des trouvailles de sépultures et de squelettes, les sépultures étant superposées l'une à l'autre. L'organisation et la nature du mobilier révèle l'identité du squelette de *Roy Mata*. Le deuxième pas de l'argumentation de Garanger est le suivant : confronter chaque énoncé du cycle mythique avec le résultat des fouilles; on assiste alors à un changement brusque de la culture matérielle. Ensuite, confronter les résultats ainsi obtenus avec les données linguistiques (voir la première partie de l'article) pour chercher éventuellement (ce qui est effectivement le cas ici) des affinités entre les langues du centre de l'archipel néo-hébridais, le micronésien et le proto-samoan; confronter finalement l'hypothèse d'une influence possible et des migrations supposées par des moyens ethnologiques proprement dits, celui des systèmes de relations inter-insulaires (systèmes de chefferies, etc.).

Sans l'étude de la littérature orale et sans le découpage des cycles mythiques en énoncés de base, Garanger le dit explicitement, l'archéologie préhistorique échouerait sans aucun doute. Mais, d'autre part, un traitement formel de ces textes ethniques révélant une organisation narrative de base, peut-être commune aux deux cycles, n'a aucun intérêt, et les chercheurs en ethno-histoire n'ont guère besoin de leçons sur la manière de découper les énoncés d'un récit.

Rappelons encore une fois l'argumentation de Garanger : les noms-de-lieux désignés par les récits sont pris comme indices pour entreprendre des fouilles archéologiques per-

mettant une datation. L'argumentation passe ensuite par quatre examens différents à savoir :

1. test 1 : géologique;
2. test 2 : découpage des énoncés;
3. test 3 : linguistique;
4. test 4 : ethnologique.

Une analyse qui se veut sémiotique n'aboutirait qu'à une organisation formelle des propriétés supposées universelles, et, par ailleurs, ne pourrait se baser que sur les énoncés, étant incapable de lier les résultats d'un découpage (analyse componentielle ou distributive) à d'autres niveaux relevant d'une autre forme d'analyse, en l'occurrence, *historique*. Nous voyons facilement, en présentant ces deux exemples d'approche génétique, qu'un traitement formel quelconque des données en question, à savoir des formes narratives en leurs différentes variantes d'une part, et des données de culture matérielle (poterie, outillage) et de sites proprement dits (sépultures, squelettes) d'autre part, s'exclue *a priori*, dans la mesure où l'*Erkenntnisinteresse* porte sur la pré- et proto-histoire des communautés sociales et de leurs représentations idéologiques.

La liste suivante donne les abréviations des titres de revues auxquelles nous nous référons :

<i>ASEMI</i>	Asie du Sud-Est et Monde Insul-indien, Paris.
<i>BSOAS</i>	Bulletin of the School of Oriental and African Studies, Londres.
<i>BEFEO</i>	Bulletin de l'École française d'extrême-orient, Hanoi-Paris.
<i>IPLS</i>	Indo-Pacific Linguistic Studies, <i>Lingua</i> , XIV et XV, 1965 (E.J.A. Henderson et G.B. Milner, eds.).
<i>JBRS</i>	Journal of the Burma Research Society, Rangoun.
<i>JSS</i>	Journal of the Siam Society, Bangkok.
<i>TBA</i>	Tribes and boundries in Australia (N. Peterson, ed.), Canberra City, Australian Institute of Aboriginal Studies, 1976.